Electricity Spot Prices Forecasting Based on Ensemble Learning

NADEELA BIBI[1], ISMAIL SHAH [ID][1], ABDELAZIZ ALSUBIE[2], SAJID ALI [ID][1], AND SHOWKAT AHMAD LONE [ID][2]

[1]Department of Statistics, Quaid-i-Azam University, Islamabad 45320, Pakistan
[2]Department of Basic Sciences, College of Science and Theoretical Studies, Saudi Electronic University, Riyadh 11673, Saudi Arabia

Corresponding authors: Ismail Shah (ishah@qau.edu.pk) and Sajid Ali (sajidali.qau@hotmail.com)

ABSTRACT Efficient modeling and forecasting of electricity prices are essential in today's competitive electricity markets. However, price forecasting is not easy due to the specific features of the electricity price series. This study examines the performance of an ensemble-based technique for forecasting short-term electricity spot prices in the Italian electricity market (IPEX). To this end, the price time series is divided into deterministic and stochastic components. The deterministic component that includes long-term trends, annual and weekly seasonality, and bank holidays, is estimated using semi-parametric techniques. On the other hand, the stochastic component considers the short-term dynamics of the price series and is estimated by time series and various machine learning algorithms. Based on three standard accuracy measures, the results indicate that the ensemble-based model outperforms the others, while the random forest and ARMA are highly competitive.

INDEX TERMS Electricity prices, forecasting, semi-parametric, IPEX, autoregressive.

I. INTRODUCTION

Before the liberalization of the electricity sector, the electrical industry was fully controlled by utility companies, generally state-owned. These utility companies were responsible for the generation, transmission, distribution, and retailing of electricity. In the early 1980s, the global electric industry underwent major changes. The monopolistic electric sector was restructured into a deregulated competitive electricity market. The Public Utility Regulatory Policies Act of 1978 (PURPA) was enacted following the energy crisis of the 1970s to encourage cogeneration and renewable resources and promote competition for electric generation [1]. Soon after, the deregulation process started in Chile in 1982 by introducing the electricity act that dissolved the state-owned monopolistic structure by commercialization and part privatization, followed by large-scale privatization in 1986. In 1990, the British electricity sector started its liberalization process, followed by Norway in 1992. Nowadays, many countries in the world have their own liberalized electricity market.

Liberalization brought important benefits to end consumers such as low prices, more choices, reliable and secure electric supply. However, modeling and forecasting of

The associate editor coordinating the review of this manuscript and approving it for publication was Vitor Monteiro [ID].

different variables related to electricity market became challenging due to the specific features introduced to this market. As electricity is a commodity that is different from other commodities in terms of its natural and physical qualities [2], electricity price forecasting is very complex due to many characteristics, including non-constant mean, high volatility, spikes or jumps, also known as extreme values [3]. In addition, electricity price prediction is generally required into three time horizons: short, medium, and long-term price prediction [4]. Price prediction for a few hours to a week is generally classified as short-term forecast (STF) which is essential for different market participants. STF helps the producers to plan the electric generation more efficiently. Other market participants require STF to develop superior bidding strategies and maximize profits in day-to-day markets [5]. A medium-term forecast (MTF) refers to a forecast made over a period ranging from a few weeks to several months. The electricity market participants require MTF for a variety of tasks, including maintenance scheduling, generation growth planning, designing investment, fuel contacting, and hedging plans [6]. The long-term price forecast (LTF) generally covers a period ranging from a few months to many years and is commonly used for site planning and selection, investment profitability analysis, and inducing power plant fuel sources [7].

The electricity prices exhibit specific characteristics that include seasonality, volatility, trend, jumps or spikes, etc. [8] and forecasting is more difficult in the presence of these unique characteristics. Due to the highly unpredictable and uncertain nature of the electricity price series, in the past, extensive studies have been made on the problem of electricity prices forecasting using different modeling techniques and procedures. To forecast electricity prices, statistical models such as regression models, time series models, and probabilistic models are commonly used. In the literature, there are a variety of time series models, such as, autoregressive (AR), moving average (MA), autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA), vector autoregressive (VAR), vector autoregression moving-average (VARMA), seasonal autoregressive integrated moving-average (SARIMA), seasonal autoregressive integrated moving-average with exogenous regressors (SARIMAX), autoregressive conditional heteroskedasticity (ARCH), and generalized autoregressive conditional heteroskedasticity (GARCH) are commonly used for a day-ahead electricity price forecasting [9]–[13]. For next-day electricity prices forecasting, [11] used a nonlinear classical time series model, GARCH, and compared the results to the ARIMA model. They evaluated the proposed methodology on Spanish and California data for hourly electricity prices and observed that the GARCH is significantly more accurate than the ARIMA. One of the most significant steps in power prediction is the shift from a deterministic view to a probabilistic view [14]–[18]. For example, to forecast next-day electricity prices, [17] proposed two modeling strategies, i.e., dynamic regression and transfer function. They used Spanish and Californian market data to evaluate the performance of their proposed methodology. For both markets, the average errors for the weeks analyzed were very low, i.e., approximately 3% and 5%, respectively. Artificial intelligence (AI) models are used to address nonlinear price forecasting issues that are not properly captured by linear time series approaches. The AI or machine learning (ML) approaches have been applied to forecast electricity prices and demand. Artificial neural networks, support vector machine, and long-short term memory (LSTM) models are widely used for electricity price forecasting problems in the literature [19], [20]. For instance, [20] proposed four different deep learning models, i.e., deep neural network (DNN), LSTM, gated recurrent unit (GRU), and convolutional neural network (CNN) for electricity price forecasting. To access the accuracy of these models, they set up a large benchmark study. The results showed that the proposed four models provide the best results with DNN is more accurate among all. In the literature, a combination of different models and algorithms are used to improve forecast accuracy. Multivariate models including functional models are used to forecast one-day-ahead electricity prices [21]–[24]. For instance, [25] used parametric and nonparametric functional models to forecast one-day-ahead electricity prices for five different electricity markets. The results indicated that functional modeling is

more efficient in predicting electricity prices than classical univariate and multivariate models. Forecast combination is often described as one of the most effective forecasting strategies. Ensemble modeling is a process in which many individual models are built to predict an outcome, either using a variety of modeling algorithms or different training data set [26]. Nowadays ensemble-based modeling framework is commonly used in the literature to improve forecast accuracy [27]–[32]. For example, [29] proposed an ensemble method which is a combination of empirical mode decomposition(EMD), support vector regression (SVR), and kernel ridge regression (KRR), for forecasting electricity prices in the short-run. Data from the Australian Energy Market Operator (AEMO) is used to access the performance of the EMD-KRR-SVR method. Six benchmark learning algorithms were used to check the accuracy of the proposed method. The results showed that the suggested approach is more efficient and accurate than the benchmark.

Although different models and techniques have been proposed in the literature for electricity price forecasting, none consistently provide accurate forecasting results. In addition, none of these works apply the ensemble approach in the context of component estimation techniques. Furthermore, in general, they do not provide any inferential analysis to test differences in the prediction accuracy of the considered models. Therefore, the main contribution of this research work is the investigation of the ensemble-based approach in the context of component estimation techniques. Furthermore, the forecasting performance is evaluated on a whole year, and the significance analysis of the differences in prediction accuracy is also investigated. In addition, the proposed model can capture the specific features of the electricity price series, leading to higher forecasting accuracy gain.

The rest of the article is organized as follows: Section II describes the methods and models used in this study. The description of the data and an application of the proposed methods and models to the Italian electricity market is given in Section III. The conclusion is given in Section IV.

II. GENERAL MODELING PROCEDURE

The main aim of this research is to forecast one-day-ahead electricity prices by using an ensemble-based modeling structure. As stated, the power price series exhibits a set of different features, and to accurately capture them in a model, the deterministic component filters the price series first, and then the residual component is simulated using a combination of time series and different machine learning algorithms. Once both the components are estimated separately, the final forecasts are obtained by combining the estimates of both deterministic and stochastic components.

Let $P_{d,j}$ denotes the price series for the dth day and j^{th} hour. Then, the price dynamics of $P_{d,j}$ can be described as follows:

$$P_{d,j} = \varphi_{d,j} + \delta_{d,j} \tag{1}$$

In other words, the price series $P_{d,j}$ is divided into two major components: $\varphi_{d,j}$ and $\delta_{d,j}$, which represent the

deterministic and stochastic components, respectively. The short-run dynamic is accounted by the component $\delta_{d,j}$. On the other hand, the long-run dynamics, different periodicities (yearly and weekly cycles), and calendar influences are all included in the $\varphi_{d,j}$ component. When they are taken into account, $\varphi_{d,j}$ is modeled as follows:

$$\varphi_{d,j} = t_{d,j} + a_{d,j} + w_{d,j} + c_{d,j} \tag{2}$$

where $t_{d,j}$ represents the long-trend, $a_{d,j}$ and $w_{d,j}$ are the annual and weekly cycles, respectively, and $c_{d,j}$ is the calendar effects. Instead of accounting for the daily periodicity in 2, each load period is modeled separately. Finally, the stochastic component $\delta_{d,j}$ (residuals) is modeled as

$$\delta_{d,j} = P_{d,j} - \hat{\varphi}_{d,j}$$
$$\delta_{d,j} = P_{d,j} - (\hat{t}_{d,j} + \hat{a}_{d,j} + \hat{w}_{d,j} + \hat{c}_{d,j}) \tag{3}$$

which is modeled by different linear and non-linear models.

A. MODELING AND FORECASTING DETERMINISTIC COMPONENT

The estimation and forecasting of the deterministic component $\varphi_{d,j}$ is explained in this section. To capture the different components described in 2, this study considers a semi-parametric approach using the generalized additive modeling technique. That is, penalized smoothing splines are used to model the long-run component (trend) $t_{d,j}$ and the yearly component $a_{d,j}$, with the series of $t_{d,j}$ and $a_{d,j}$ considered as a functional objects of the time and yearly cycle. Dummy variables are used to describe weekly cycle $w_{d,j}$ and bank holidays $c_{d,j}$. In the case of weekly periodicity, seven dummy variables are required, i.e., $w_{d,j} = \sum_{i=1}^{7} \gamma_i I_{d,j}$ where $I_{d,j} = 1$ if d is the ith day of the week, and 0 otherwise. In the case of bank holidays, only two dummy variables are used, i.e., $c_{d,j} = \sum_{k=1}^{2} \alpha_k I_{d,j}$, where $I_{d,j} = 1$ if d indicates a bank holiday and 0 otherwise. The parameters γ_i and α_k are estimated using the ordinary least squares (OLS) methods. Once these components are estimated, their one-day-ahead forecasts are straightforward as all these components are the deterministic functions of time or calendar conditions.

B. MODELING AND FORECASTING STOCHASTIC COMPONENT

Following the estimation of the deterministic components using a semi-parametric approach, the stochastic (residual) component $\delta_{d,j}$ is modeled using a combination of autoregressive moving average (ARMA) and different machine learning models, such as neural network auto-regressive (NAR), random forest (RF), support vector regression (SVR), gradient boosting machine (GBM), and an ensemble model based on these methods. Once these models are estimated and a one-day-ahead forecast is obtained for each model, the final one-day-ahead out-of-sample forecast is obtained by combining both the deterministic and stochastic components as follows.

$$\hat{P}_{d+1,j} = (\hat{t}_{d+1,j} + \hat{a}_{d+1,j} + \hat{w}_{d+1,j} + \hat{c}_{d+1,j} + \hat{\delta}_{d+1,j})$$

$$\hat{P}_{d+1,j} = \hat{\varphi}_{d+1,j} + \hat{\delta}_{d+1,j} \tag{4}$$

The details about each model used for the estimation of the stochastic component is given in the following.

1) AUTOREGRESSIVE MOVING AVERAGE (ARMA) MODELS

A powerful tool for modeling univariate time series is an autoregressive moving average (ARMA) model. In general, an ARMA model is a combination of autoregressive (AR) and moving average (MA) models. The mathematical form for an ARMA of order (s, r), abbreviated as ARMA(s, r), is as follows.

$$\delta_{d,j} = \beta + (\sum_{i=1}^{s} \psi_i \delta_{d-i,j}) + (\sum_{k=0}^{r} \vartheta_k \zeta_{d-k,j}) \tag{5}$$

where β represents the constant term (intercept), ψ_j ($j = 1, 2, \cdots, s$) and ϑ_i ($i = 0, 1, 2, \cdots, r$) are the AR and MA parameters, respectively, and $\zeta_{d-r,j}$ is a white noise series with zero mean and variance σ_ζ^2. In general, different information criteria or auto-correlation function (ACF) and partial auto-correlation function (PACF) are used to specify the order of s and r.

2) ARTIFICIAL NEURAL NETWORK AUTOREGRESSIVE (NNAR)

Generally, a neural network is a network or circuit of neurons. An artificial neural network is made up of artificial neurons or nodes. Modeling complex nonlinear relationships between the response variable and its predictors are possible using neural network models. A feedback neural network is assembled with delayed time series values as input and a hidden layer with dimension nodes. There are at least three layers of nodes in an autoregressive neural network (NNAR): an input layer, a hidden layer, and an output layer. The outputs of one layer are used as inputs to the next. The nonlinear autoregressive neural network can be trained to forecast a time series $\delta_{d,j}$ given its previous values $\delta_{d-1,j}, \delta_{d-2,j}, \cdots, \delta_{d-p,j}$, which are referred as feedback delays, with d is the time delay parameter. The term NNAR(p,k) implies that the hidden layer contains p delayed inputs and k nodes. An NNAR(p,0) is the same as an ARMA(p,0), but it does not include the limitations of the parameters that ensure stationarity. A nonlinear autoregressive neural network used for time series prediction can be described as follows.

$$\delta_{d,j} = \gamma(\delta_{d-1,j}, \delta_{d-2,j}, \ldots, \delta_{d-p,j}) + \zeta_{d,j} \tag{6}$$

Since the $\gamma(.)$ function for the next day is unknown, the neural network training aims to approximate it by optimizing the network weights and neural bias. Consequently, the following equation precisely specifies an NNAR model.

$$\delta_{d,j} = \alpha + \sum_{i=1}^{p} \alpha_i \theta(\sum_{k=1}^{q} \phi_{k_i} \delta_{d-k,j} + \phi_i) + \zeta_{d,j} \tag{7}$$

where q is the number of entries, p is the number of hidden layers with activation function θ, and ϕ_{k_i} is the parameter corresponding to the weight of the connection between the
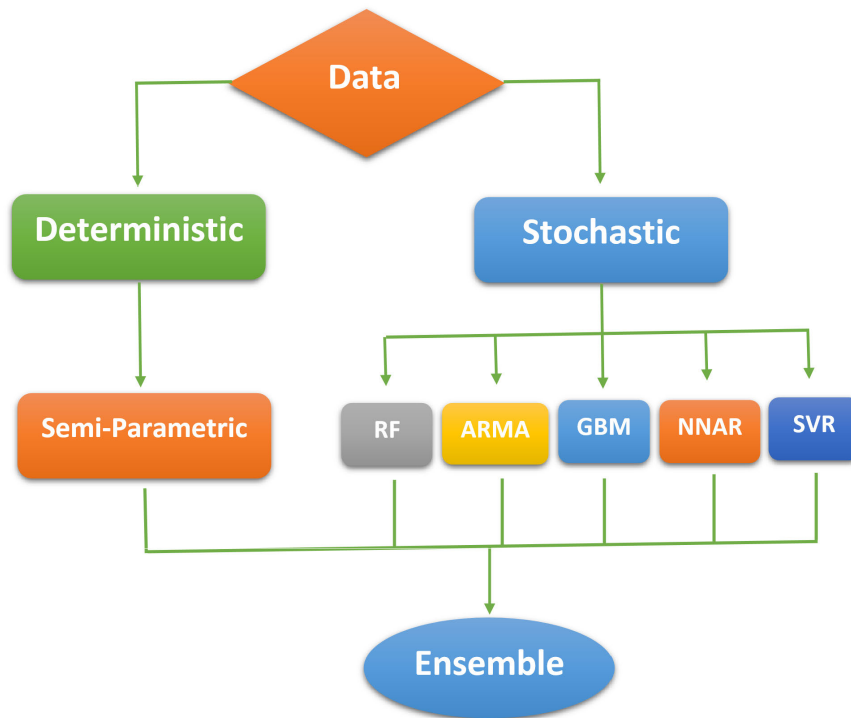
FIGURE 1. Flowchart of the proposed modeling framework.

input unit k and the hidden unit i, α_i is the weight of the connection between the hidden unit i and the output unit, and ϕ_i and α are the constants that correspond to the hidden unit i and the output unit, respectively. In this work, NNAR(2,3) is used which implies that the hidden layer contains 2 delayed inputs and 3 nodes.

3) RANDOM FOREST (RF)

Random decision forests, often known as random forests, are ensemble learning methods for classification and regression that work by training a large number of decision trees. Breiman *et al.* [33] presented the procedures of random forest (RF), classification and regression tree (CART). In 1996, Breiman proposed another significant RF technique called Bagging [34]. The RF is a well-known machine learning algorithm that falls within the supervised learning category. The RF algorithm is applied in a wide range of areas, including finance, stock trading, health care, and e-commerce. It produces a forest out of a collection of decision trees that are often trained using the bagging method. The RF algorithm determines the outcome based on the decision trees' predictions. It forecasts by averaging the output of various trees. The precision of the result improves as the number of trees grows. The higher the number of trees in the forest, the more accurate it is and the problem of over-fitting is avoided.

The random forest training algorithm uses the common technique of bootstrap aggregation, or bagging, to train tree learners. Bagging repeatedly (M times) takes a random sample with replacement of the training set and fits trees to these samples given a training set $Y = y_1, \ldots, y_n$ with responses $X = x_1, \ldots, x_n$. For $m = 1, \ldots, M$, the algorithm works as follows.

1. Sample m training data points from X and Y with replacement and denote them Y_m and X_m.

2. Train a classification or regression tree f_m on Y_m and X_m.

After training, summing the predictions from all the various regression trees on y' can be used to make predictions for unseen samples y'.

$$\hat{f} = \frac{1}{M} \sum_{m=1}^{M} f_m(y') \tag{8}$$

By reducing the variance of the model without increasing the bias, this bootstrap approach improves the performance of the model. In the case of classification trees, the majority vote is used.

4) GRADIENT BOOST MACHINES (GBM)

The gradient boost machine (GBM) is a type of ensemble learning that uses a sequential learning process to create a good classification or regression model. Initially, the data are fitted to a regression tree, and predictions and the initial residual are produced based on this information. A new model is fitted to the previous residual, followed by a new forecast, which includes the initial forecast and finally a new residual. This procedure is continued iteratively until a convergence condition is satisfied. At each iteration, a new model is fitted to the data to compensate for the deficiencies of the prior model.

The GBM algorithm works best when the contribution of the additional decision tree is minimized at each step of the iteration by using shrinkage parameter β, also known as the learning rate. In the GBM, the idea behind the shrinkage procedure is that more small passes provide greater precision than fewer large passes. The learning parameter β has a range of values between 0 and 1; the smaller the value, the more accurate the model. However, because the value of β is inversely proportional to the number of iterations, choosing a stronger shrinkage (smaller β) implies a higher number of iterations to achieve convergence. The following steps provide a simplified illustration of this algorithm:

1. Sets the hyper parameters of GBM: depth of the decision trees τ, the number of iterations p, the shrinkage parameter β, and the fraction of the subsample, n.

2. Set $\zeta_0 = x$ and $f = 0$ for the residual. The mean value of x has also been proposed as an initial estimate of f.

3. For p = 1,2,...,P, follow:

a. Select a subsample $\{xi, yi\}^{N^*}$ at random from the entire training dataset, where N^* is the number of data points corresponding to the fraction n.

b. Fit a decision tree f^p of depth τ to the residual ζ_{p-1} value using $\{xi, yi\}^{N^*}$.

c. Add the decision tree to the model to update f.

$$f(y) \leftarrow f(y) + \beta f^p(y) \qquad (9)$$

d. Update the residual

$$\zeta_p \leftarrow \zeta_{p-1} - \beta f^p(Y) \qquad (10)$$

There are four hyperparameters in the GBM: (1) decision tree depth ($\tau = 1$), which also affects the model's maximum interaction order, (2) number of iterations (p), which is same as the number of decision trees, (3) learning rate ($\beta = 0.001$), which is normally a small positive value between 0 and 1, (4) fraction of the subsample (n), which is the percentage of data used in each iteration step.

5) SUPPORT VECTOR REGRESSION (SVR)

Support vector machine can also be applied as a regression approach while retaining all of the algorithm's key properties (maximal margin) [35]. With a few minor exceptions, the support vector regression (SVR) uses the same classification concepts as the SVM. The SVR involves finding support vectors adjacent to a hyperplane that optimizes the gap between two-point classes based on the difference between the target and a threshold. The SVR uses kernel functions to calculate the similarity between two observations to handle non-linear situations. The capability of the SVR to capture predictor non-linearity and then use it to improve forecasting situations is one of the key advantages of using it. The SVR helps to determine how much error is acceptable in the model. Support vector regression equation with kernel function can be written as

$$\delta_{t,j} = \sum_{i=1}^{P} (\beta_i - \beta_i^*) K(y_i, y) + \alpha \qquad (11)$$

The kernel function, $K(y_i, y)$, represents the inner product, while α is accommodated within the kernel function.

$$K(y_i, y_k) = \exp\left(-\frac{\|y_i - y_k\|^2}{2\sigma^2}\right) \qquad (12)$$

where $\|y_i - y_k\|^2$ is the Euclidean distance between the two feature vectors squared.

6) ENSEMBLE MODEL

Ensemble methods are strategies that combine multiple models to achieve better results. In most cases, ensemble approaches provide more accurate results than a single model. In many machine learning competitions, the winning solutions used ensemble approaches. In this work, five models are used to make an ensemble namely, autoregressive moving average (ARMA), neural network auto-regressive (NAR), random forest (RF), support vector regression (SVR), and gradient boosting machine (GBM). To make an ensemble model weighting scheme is applied. A weighted average ensemble is a method for allowing different models to contribute to a forecast in proportion to their level of confidence or estimated performance. The contribution of each member to the final forecast is weighted by the model's performance in a weighted ensemble, which is an extension of a model averaging ensemble. The model weights are small positive values and the sum of all weights equals one, indicating the percentage of trust or expected performance from each model.

7) NAÏVE MODEL

This study used the benchmark model proposed by [36]. Although it is a naïve model, the forecasting accuracy of this approach is much better when compared with other naïve models. This method forecasts a day based on the information contained in all the preceding days. For example, if we are interested in predicting a Saturday, we select the day before Saturday, i.e., Friday. We then compare this Friday with all available Fridays in the training set and calculate their errors. Based on the error, we decide which Friday is more similar to the current Friday. Once a similar Friday is identified, we select the next day to the selected Friday, i.e., Saturday, and use it as a forecast for the Saturday we are interested in. We do the same procedure for all days of the week and for the whole year.

III. OUT-OF-SAMPLE PRICE FORECASTING

The electricity price data from the Italian electricity market (IPEX) is used in this work to evaluate the forecasting performance of different models. The IPEX is an important electricity market that is fully liberalized since 2007. For this study, the data includes 52584 observations over 2191 days, ranges from January 1st, 2014 to December 31st, 2019. The data collection contains 24 observations for each day corresponds to a load period and the prices are given in Euro/MWh. The descriptive statistics for the data under study are given
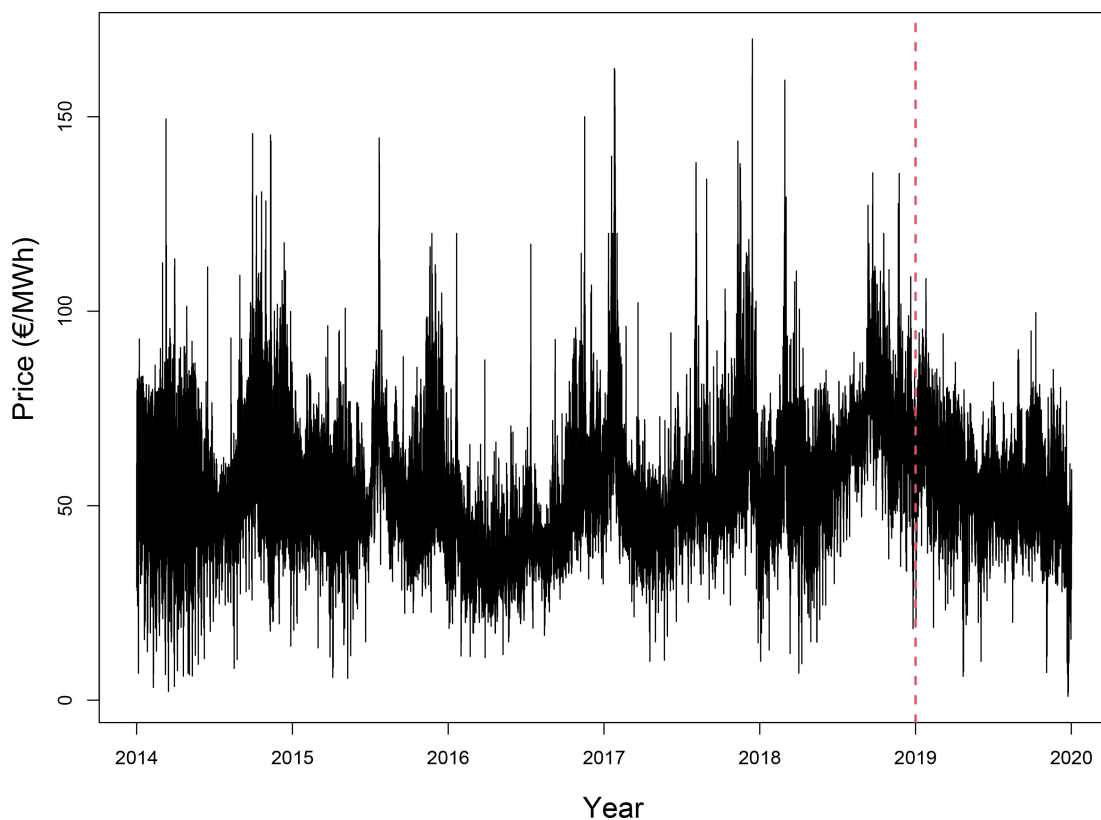
FIGURE 2. IPEX: Electricity price time series from January 1st, 2014 to December 31st, 2019.

TABLE 1. IPEX: Descriptive statistics for electricity price time series for the period January 1st, 2014 to December 31st, 2019.

	Min	Max	Mean	SD
Monday	7.51	176.61	58.14	20.03
Tuesday	18.74	324.20	58.70	19.26
Wednesday	0.10	205.39	59.34	20.08
Thursday	5.89	212.84	59.06	20.10
Friday	14.24	222.25	58.44	18.91
Saturday	9.13	147.84	53.86	17.65
Sunday	0.10	146.67	48.70	19.02
Total	0.10	324.20	56.60	19.65

TABLE 2. Electricity Prices: Mean absolute error (MAE), root mean square error (RMSE) and Person correlation coefficient (R) for one-day-ahead out-of-sample forecast.

Models	MAE	RMSE	R
RF	5.225049	6.842467	0.8414007
ARMA	5.184236	6.785045	0.8449486
GBM	5.306283	6.967804	0.8361897
NNAR	5.293977	6.970038	0.8533544
SVR	5.393675	7.085886	0.8315566
Ensemble	5.049538	6.664433	0.8516408
Naïve	6.664433	9.250304	0.7436619

in Table 1. The table provides the minimum, the maximum, and the mean prices as well as the standard deviation of the weekdays' prices. The table shows that the minimum price reaches 0.10 for Wednesday and Sunday while the maximum price is 324.20 for Tuesday. The mean prices for the weekdays range from 48.70 to 59.34. The last column provides the standard deviation values for each day of the week, while the last row describes the descriptive statistics for the whole electricity price time series. For modeling and forecasting, the data set is divided into two groups: For model identification and estimation, we used data from January 1st, 2014 to December 31st, 2018 (43824 hourly observations/ 1826 days), and January 1st, 2019 to December 31st, 2019

(8760 hourly observations/ 365 days) for analyzing the models' one-day-ahead out-of-sample forecasting accuracy using the expending window technique. The electricity price time series is plotted in Figure 2 where the vertical red dotted line divided the estimation and out-of-sample forecasting

TABLE 3. P-values for the DM test. Null hypothesis: equal prediction accuracy, alternative hypothesis: model in the row is more accurate than model in the column (squared loss function used).

Models	RF	ARMA	GBM	SVR	NNAR	Ensemble	Naïve
RF	-	>0.99	<0.01	<0.01	0.17	>0.99	<0.01
ARMA	<0.01	-	<0.01	<0.01	0.03	>0.99	<0.01
GBM	>0.99	>0.99	-	<0.01	0.82	>0.99	<0.01
SVR	>0.99	>0.99	>0.99	-	>0.99	>0.99	<0.01
NNAR	0.82	<0.01	0.17	<0.01	-	>0.99	<0.01
Ensemble	<0.01	<0.01	<0.01	<0.01	<0.01	-	<0.01
Naïve	>0.99	>0.99	>0.99	>0.99	>0.99	>0.99	-

periods. The main reason for using a whole year as an out-of-sample forecasting period is to assess the performance of each model in every type of situation during a year. Since the prices vary throughout the year, with some periods producing extreme prices, evaluating models' performance on a complete year is a more realistic approach. The deterministic component is first estimated and forecasted to obtain the stochastic (residual) part. For the stochastic component, all the models are used to estimate and forecast the stochastic component. This study considers three standard accuracy measures, namely mean absolute error (MAE), root mean squared error (RMSE), and Pearson correlation (R), to assess the model's forecasting performance. Mathematically, they are defined as

$$\text{MAE} = \text{mean}(|P_{d,j} - \hat{P}_{d,j}|)$$
$$\text{RMSE} = \sqrt{\text{mean}(P_{d,j} - \hat{P}_{d,j})^2}$$
$$\text{R} = \text{corr}(P_{d,j}, \hat{P}_{d,j}).$$

The observed and forecasted prices for the dth day (d = 1, 2, . . . , 365) and the jth load period (j = 1, 2, . . . , 24) are $P_{d,j}$ and $\hat{P}_{d,j}$, respectively.

Table 2 shows the out-of-sample one-day-ahead forecast findings for the entire year 2019. The models used to estimate the stochastic component are listed in the first column. The MAE, RMSE, and R are the accuracy measures listed in the second, third, and fourth columns, respectively. From the table, it is clear that the ensemble-based model outperforms the individual models in terms of forecasting. The ensemble model produced 5.05, 6.66, and 0.85 values for the MAE, MSE, and R, respectively. Second, the ARMA with MAE 5.18 and RF with MAE 5.23 achieved slightly better results suggest that they are strong competitors. The results of the NNAR with MAE 5.29, GBM with MAE 5.30, and SVR with MAE 5.39 are also good, but not satisfactory as compared to

the MAE of the ensemble model. Moreover, the forecasting accuracy of the naïve model is much lower than all other models used in the study. This shows that the current modeling framework is efficient in predicting one-day-ahead electricity prices.

Several methods are available in the literature to compare the performance of different forecasting models [37]. In this work, the Diebold and Mariano (DM) [38] test is used to determine the significance of the differences between the results reported in Table 2, and the results are listed in Table 3. The null hypothesis of the DM test corresponds to equal forecast accuracy, while the alternative hypothesis states that the model in the row is more accurate than in the column. The DM results listed in Table 3 indicate that the proposed ensemble model is statistically significant than all other models. In addition, the proposed modeling structure is also statistically significant than the naïve model.

Finally, the obtained results are compared with the results provided in the literature and found to be highly comparative. It is worth mentioning that such comparisons are very difficult as different authors use different indicators, forecasting periods, and forecasting horizons. For example, the work of [39] reported an MAE of 8.58, whereas we obtained a value of 5.05 with the ensemble model. Ref. [40] reported an RMSE of 11.58 obtained with an ARX-EGARCH model, whereas we obtained 6.66 with the ensemble model. Ref. [41] obtained an RMSE of 16.72 and 15.79 using ARMA and GARCH models, respectively, which is considerably higher than our value of 6.66.

IV. CONCLUSION

The forecasting problem of electricity prices in the recently liberalized market is analyzed in depth in this work. In competitive electricity markets, efficient modeling and forecasting of electricity prices are essential as forecasts are necessary

for risk management, trading, and future planning because prices and demand are determined one day before physical delivery. However, price forecasting is not easy as electricity prices can have complex characteristics, such as high volatility, seasonality, calendar impact, non-stationarity, non-linearity, mean reversion, etc. Thus, this research work aims to propose an ensemble-based technique to improve prediction accuracy by combining time series and different machine learning algorithms. To this end, the deterministic component filters the price series first, and then the residual component is modeled using an ensemble-based method to accurately capture the specific properties of electricity prices in the model. The deterministic component that includes long-term trends, annual and weekly seasonality, and bank holidays, is estimated using a semi-parametric approach. The stochastic component, on the other hand, considers the short-term dynamics of the price series and is modeled by the autoregressive moving average (ARMA), and various machine learning algorithms, including neural network autoregressive (NNAR), random forest (RF), support vector regression (SVR), gradient boosting machine (GBM). After modeling the deterministic and stochastic components individually, the final prediction is obtained by combining the estimates of both components. The proposed modeling approach is presented together with an application to the Italian electricity market (IPEX). To evaluate the performance of our models, data from the period January 1st, 2014 to December 31st, 2019 is used, whereas from January 1st, 2014 to December 31st, 2018 is used for model estimation while the entire 2019 year is for the one-day-ahead out-of-sample forecast. The mean absolute error (MAE), root mean square error (RMSE), and Pearson's correlation coefficient (r) are used to assess the effectiveness of the ensemble-based methodology.

The results suggest that the proposed ensemble method is efficient in predicting electricity prices. The ensemble model produced the MAE value of 5.05, whereas the ARMA and RF produced 5.18 and 5.23, respectively. The results of the NNAR with MAE 5.29, GBM with MAE 5.30, and SVR with MAE 5.39 are also good, but not satisfactory as compared to the MAE of the ensemble model. Moreover, the forecasting accuracy of the naïve model with the MAE of 6.66 is much lower than all other models used in the study. The DM test validated the significance of the ensemble model compared to the rest. Lastly, as the ensemble-based model outperforms the others, the RF and ARMA produce better results than other models used in this study. In conclusion, the proposed modeling framework is statistically significant in forecasting electricity prices. As the current study considers only the IPEX, in the future, it can be extended to other markets to assess the performance of the proposed models.

REFERENCES
[1] P. L. Joskow, "Lessons learned from electricity market liberalization," *Energy J.*, vol. 29, no. 1, pp. 9–42, Sep. 2008.

[2] R. J. Serrallés, "Electric energy restructuring in the European union: Integration, subsidiarity and the challenge of harmonization," *Energy Policy*, vol. 34, no. 16, pp. 2542–2551, Nov. 2006.

[3] A. J. Conejo, M. A. Plazas, R. Espinola, and A. B. Molina, "Day-ahead electricity price forecasting using the wavelet transform and ARIMA models," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 1035–1042, May 2005.

[4] I. Shah, H. Iftikhar, and S. Ali, "Modeling and forecasting medium-term electricity consumption using component estimation technique," *Forecasting*, vol. 2, no. 2, pp. 163–179, May 2020.

[5] G. Li, C. C. Liu, C. Mattson, and J. Lawarree, "Day-ahead electricity price forecasting in a grid environment," *IEEE Trans. Power Syst.*, vol. 22, no. 1, pp. 266–274, Feb. 2007.

[6] S. S. Torbaghan, A. Motamedi, H. Zareipour, and L. A. Tuan, "Medium-term electricity price forecasting," in *Proc. North Amer. Power Symp. (NAPS)*, Sep. 2012, pp. 1–8.

[7] M. Cerjan, A. Petričić, and M. Delimar, "HIRA model for short-term electricity price forecasting," *Energies*, vol. 12, no. 3, p. 568, Feb. 2019.

[8] R. Weron, *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*, vol. 403. Hoboken, NJ, USA: Wiley, 2007.

[9] I. Shah, H. Iftikhar, S. Ali, and D. Wang, "Short-term electricity demand forecasting using components estimation technique," *Energies*, vol. 12, no. 13, p. 2532, Jul. 2019.

[10] T. Jakasa, I. Androcec, and P. Sprcic, "Electricity price forecasting—ARIMA model approach," in *Proc. 8th Int. Conf. Eur. Energy Market (EEM)*, May 2011, pp. 222–225.

[11] R. C. Garcia, J. Contreras, M. V. Akkeren, and J. B. C. Garcia, "A GARCH forecasting model to predict day-ahead electricity prices," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 867–874, May 2005.

[12] Z. Tan, J. Zhang, J. Wang, and J. Xu, "Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models," *Appl. Energy*, vol. 87, no. 11, pp. 3606–3610, Nov. 2010.

[13] F. Lisi and I. Shah, "Forecasting next-day electricity demand and prices based on functional models," *Energy Syst.*, vol. 11, no. 4, pp. 947–979, Nov. 2020.

[14] T. Hong, P. Pinson, Y. Wang, R. Weron, D. Yang, and H. Zareipour, "Energy forecasting: A review and outlook," *IEEE Open Access J. Power Energy*, vol. 7, pp. 376–388, 2020.

[15] Q. Chang, Y. Wang, X. Lu, D. Shi, H. Li, J. Duan, and Z. Wang, "Probabilistic load forecasting via point forecast feature integration," in *Proc. IEEE Innov. Smart Grid Technol. Asia (ISGT Asia)*, May 2019, pp. 99–104.

[16] T. Serafin, B. Uniejewski, and R. Weron, "Averaging predictive distributions across calibration Windows for day-ahead electricity price forecasting," *Energies*, vol. 12, no. 13, p. 2561, Jul. 2019.

[17] F. J. Nogales, J. Contreras, A. J. Conejo, and R. Espinola, "Forecasting next-day electricity prices by time series models," *IEEE Trans. Power Syst.*, vol. 17, no. 2, pp. 342–348, May 2002.

[18] I. Shah, S. Akbar, T. Saba, S. Ali, and A. Rehman, "Short-term forecasting for the electricity spot prices with extreme values treatment," *IEEE Access*, vol. 9, pp. 105451–105462, 2021.

[19] Y. Wang, Q. Chen, N. Zhang, and Y. Wang, "Conditional residual modeling for probabilistic load forecasting," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 7327–7330, Nov. 2018.

[20] J. Lago, F. De Ridder, and B. De Schutter, "Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms," *Appl. Energy*, vol. 221, pp. 386–405, Jul. 2018.

[21] I. Shah and F. Lisi, "Forecasting of electricity price through a functional prediction of sale and purchase curves," *J. Forecasting*, vol. 39, no. 2, pp. 242–259, Mar. 2020.

[22] J. M. Vilar, R. Cao, and G. Aneiros, "Forecasting next-day electricity demand and price using nonparametric functional methods," *Int. J. Elect. Power Energy Syst.*, vol. 39, no. 1, pp. 48–55, 2012.

[23] I. Shah and F. Lisi, "Day-ahead electricity demand forecasting with nonparametric functional models," in *Proc. 12th Int. Conf. Eur. Energy Market (EEM)*, May 2015, pp. 1–5.

[24] Y. Chen and B. Li, "An adaptive functional autoregressive forecast model to predict electricity price curves," *J. Bus. Econ. Stat.*, vol. 35, no. 3, pp. 371–388, 2017.

[25] I. Shah, "Modeling and forecasting electricity market variables," Ph.D. dissertation, Dept. Stat. Sci., Univ. Padova, Padua, Italy, 2016.

[26] V. Kotu and B. Deshpande, *Predictive Analytics and Data Mining: Concepts and Practice With Rapidminer*. San Mateo, CA, USA: Morgan Kaufmann, 2014.

[27] B. Neupane, W. L. Woon, and Z. Aung, "Ensemble prediction model with expert selection for electricity price forecasting," *Energies*, vol. 10, no. 1, p. 77, Jan. 2017.

[28] R. K. Agrawal, F. Muchahary, and M. M. Tripathi, "Ensemble of relevance vector machines and boosted trees for electricity price forecasting," *Appl. Energy*, vol. 250, pp. 540–548, Sep. 2019.

[29] X. Qiu, P. N. Suganthan, and G. A. J. Amaratunga, "Short-term electricity price forecasting with empirical mode decomposition based ensemble kernel machines," *Proc. Comput. Sci.*, vol. 108, pp. 1308–1317, Jan. 2017.

[30] S. Barak and S. S. Sadegh, "Forecasting energy consumption using ensemble ARIMA–ANFIS hybrid algorithm," *Int. J. Elect. Power Energy Syst.*, vol. 82, pp. 92–104, Nov. 2016.

[31] M. Ribeiro, S. Stefenon, J. de Lima, A. Nied, V. Mariani, and L. Coelho, "Electricity price forecasting based on self-adaptive decomposition and heterogeneous ensemble learning," *Energies*, vol. 13, no. 19, p. 5190, Oct. 2020.

[32] M. Abdel-Nasser, K. Mahmoud, and M. Lehtonen, "Reliable solar irradiance forecasting approach based on choquet integral and deep LSTMs," *IEEE Trans. Ind. Informat.*, vol. 17, no. 3, pp. 1873–1881, Mar. 2021.

[33] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. New York, NY, USA: Chapman & Hall, 1984.

[34] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, 1996.

[35] B. T. Ojemakinde, "Support vector regression for non-stationary time series," M.S. thesis, Graduate School, Univ. Tennessee, Knoxville, TN, USA, 2006.

[36] I. Shah, H. Bibi, S. Ali, L. Wang, and Z. Yue, "Forecasting one-day-ahead electricity prices for Italian electricity market using parametric and nonparametric approaches," *IEEE Access*, vol. 8, pp. 123104–123113, 2020.

[37] N. D. Bokde, Z. M. Yaseen, and G. B. Andersen, "ForecastTB—An R package as a test-bench for time series forecasting—Application of wind speed and solar radiation modeling," *Energies*, vol. 13, no. 10, p. 2578, May 2020.

[38] F. X. Diebold and R. S. Mariano, "Comparing predictive accuracy," *J. Bus. Econ. Stat.*, vol. 13, no. 3, pp. 253–263, 1995.

[39] A. Petrella and S. Sapio, "No PUN intended: A time series analysis of the Italian day-ahead electricity prices," Florence School Regulation, Loyola de Palacio Programme Energy Policy, EUI RSCAS, Tech. Rep., Mar. 2010. [Online]. Available: http://hdl.handle.net/1814/13099

[40] A. Petrella and S. Sapio, "A time series analysis of day-ahead prices on the Italian power exchange," in *Proc. 6th Int. Conf. Eur. Energy Market*, May 2009, pp. 1–6.

[41] A. Cervone, E. Santini, S. Teodori, and D. Z. Romito, "Electricity price forecast: A comparison of different models to evaluate the single national price in the Italian energy exchange market," *Int. J. Energy Econ. Policy*, vol. 4, no. 4, pp. 744–758, 2014.

NADEELA BIBI is currently pursuing the M.Phil. degree in statistics with the Department of Statistics, Quaid-i-Azam University (QAU), Islamabad, Pakistan. Her research interests include time series analysis, forecasting electricity market variables, and applied statistics.

ISMAIL SHAH received the master's degree from Lund University, Sweden, and the Ph.D. degree from the University of Padua, Italy. He is currently working as an Assistant Professor with the Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan. His research interests include functional data analysis, time series analysis, regression analysis, energy economics, and applied and industrial statistics. He is also working as an Editor of the *Journal of Quantitative Methods*.
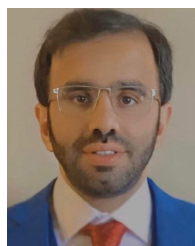
ABDELAZIZ ALSUBIE received the Ph.D. degree in statistics from North Dakota State University, in 2020, under the supervision of Prof. Rhonda Magel. He is currently working as an Assistant Professor of statistics with the Department of Basic Sciences, Saudi Electronic University. His research interests include nonparametric, inference under order restriction, and simulation studies.

SAJID ALI received the Ph.D. degree in statistics from Bocconi University, Milan, Italy. He is currently an Assistant Professor with the Department of Statistics, Quaid-i-Azam University (QAU), Islamabad, Pakistan. His research interests include time series analysis, Bayesian inference, construction of new flexible probability distributions, and process monitoring.

SHOWKAT AHMAD LONE received the M.Sc. degree in statistics from the University of Kashmir, in 2013, and the Ph.D. degree in statistics from Aligarh Muslim University, India, in 2017. For more than seven years, he has been teaching courses in statistics, mathematics, management science, economics, and other quantitative models. He is currently an Assistant Professor of mathematics and statistics with Saudi Electronic University (SEU), Saudi Arabia. He has authored textbook chapters in reliability engineering and statistics. His research interests include the area of reliability engineering, computer science, and related fields. He has published research articles in reputed international journals in the fields of mathematical and statistical sciences. He is a referee for several international journals in the field of pure and applied statistics.

● ● ●